**WEEK 5: Streamlined Energy Analytics and Prediction System with Azure**

**Prerequisites:**

1. Set up an Azure DevOps account and project.
2. Access Azure Databricks workspace for notebooks and clusters.
3. Generate a Personal Access Token (PAT) in Databricks.
4. Install Databricks CLI on your local machine.

**Step 1: Set up Databricks CLI**

1. **Install Databricks CLI:**

pip install databricks-cli

1. Configure the CLI:

databricks configure –token

**Step 2: Create an Azure DevOps Pipeline**

1. Create a new YAML pipeline in Azure DevOps.
2. Add variables for:
   * DATABRICKS\_HOST: Your Databricks URL.
   * DATABRICKS\_TOKEN: The generated token.

**Step 3: Example Pipeline YAML**

trigger:

- main

pool:

vmImage: 'ubuntu-latest'

variables:

DATABRICKS\_HOST: 'https://<databricks-instance>.azuredatabricks.net'

DATABRICKS\_TOKEN: $(databricksToken)

steps:

# Install Python and Databricks CLI

- task: UsePythonVersion@0

inputs:

versionSpec: '3.x'

addToPath: true

- script: |

pip install databricks-cli

displayName: 'Install Databricks CLI'

# Configure Databricks CLI

- script: |

databricks configure --host $(DATABRICKS\_HOST) --token $(DATABRICKS\_TOKEN)

displayName: 'Configure Databricks CLI'

# Upload Notebook to Databricks

- script: |

databricks workspace import ./notebooks/Energy\_Monitor.py /Shared/Energy\_Monitor\_Notebook -l PYTHON

displayName: 'Upload Notebook to Databricks'

# Run Databricks Notebook

- script: |

JOB\_ID=$(databricks runs submit --json-file notebook\_run\_config.json | jq -r '.run\_id')

echo "Job ID: $JOB\_ID"

databricks runs wait --run-id $JOB\_ID

displayName: 'Run Databricks Notebook'

**Step 4: JSON Config File** (notebook\_run\_config.json)

{

"run\_name": "Energy Monitoring Run",

"new\_cluster": {

"spark\_version": "10.4.x-scala2.12",

"node\_type\_id": "Standard\_DS3\_v2",

"num\_workers": 2

},

"notebook\_task": {

"notebook\_path": "/Shared/Energy\_Monitor\_Notebook",

"base\_parameters": {

"input\_param1": "valueA",

"input\_param2": "valueB"

}

}

}

**Key Points:**

* **Databricks CLI**: Upload and run notebooks.
* **Azure DevOps Variables**: Keep tokens secure in variables.
* **Run Configuration**: Customize the JSON for cluster and notebook settings.